

OLPE  
TEGEC CENTER 1600/2900  
09/13/2001  
STIC sub

RECEIVED

Serial Number: 09/471,255

ENTERED

- ☐ Changed a file from non-ASCII to ASCII
- ☐ Changed the margins in cases where the sequence text was "wrapped" down to the next line.
- ☐ Edited a format error in the Current Application Data section, specifically: \_\_\_\_\_
- ☐ Edited the Current Application Data section with the actual current number. The number inputted by the applicant was ☐ the prior application data; or ☐ other \_\_\_\_\_
- ☐ Added the mandatory heading and subheadings for "Current Application Data".
- ☐ Edited the "Number of Sequences" field. The applicant spelled out a number instead of using an integer.
- ☐ Changed the spelling of a mandatory field (the headings or subheadings), specifically: \_\_\_\_\_
- ☐ Corrected the SEQ ID NO when obviously incorrect. The sequence numbers that were edited were: \_\_\_\_\_
- ☐ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited: \_\_\_\_\_
- ☐ Corrected subheading placement. All responses must be on the same line as each subheading. If the applicant placed a response below the subheading, this was moved to its appropriate place.
- ☐ Inserted colons after headings/subheadings. Headings edited included: \_\_\_\_\_
- ☐ Deleted extra, invalid, headings used by an applicant, specifically: \_\_\_\_\_
- ☒ Deleted: ☒ non-ASCII "garbage" at the beginning/end of files; ☐ secretary initials/filename at end of file; ☐ page numbers throughout text; ☐ other invalid text, such as \_\_\_\_\_
- ☐ Inserted mandatory headings, specifically: \_\_\_\_\_
- ☐ Corrected an obvious error in the response, specifically: \_\_\_\_\_
- ☐ Edited identifiers where upper case is used but lower case is required, or vice versa.
- ☐ Corrected an error in the Number of Sequences field, specifically: \_\_\_\_\_
- ☐ A "Hard Page Break" code was inserted by the applicant. All occurrences had to be deleted.
- ☐ Deleted *ending* stop codon in amino acid sequences and adjusted the "(A)Length:" field accordingly (error due to a PatentIn bug). Sequences corrected: \_\_\_\_\_
- ☐ Other: \_\_\_\_\_

Examiner: The above corrections must be communicated to the applicant in the first Office Action. DO NOT send a copy of this form.

## RAW SEQUENCE LISTING

DATE: 10/09/2001

PATENT APPLICATION: US/09/471,255

TIME: 18:08:44

Input Set : A:\PTO.amc

Output Set: N:\CRF3\10092001\I471255.raw

```

4 <110> APPLICANT: BIOCHEM PHARMA INC.
5   HAMEL, Jose
6   BRODEUR, Bernard R.
7   PINEAU, Isabelle
8   MARTIN, Denis
9   RIOUX, Clment
10  CHARLAND, Nathalie
12 <120> TITLE OF INVENTION: NOVEL STREPTOCOCCUS ANTIGENS
15 <130> FILE REFERENCE: 12806-11PCT
C--> 17 <140> CURRENT APPLICATION NUMBER: US/09/471,255
C--> 17 <141> CURRENT FILING DATE: 1999-12-23
17 <150> PRIOR APPLICATION NUMBER: US 60/113,800
18 <151> PRIOR FILING DATE: 1998-12-23
20 <160> NUMBER OF SEQ ID NOS: 102
22 <170> SOFTWARE: FastSEQ for Windows Version 3.0
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 3120
26 <212> TYPE: DNA
27 <213> ORGANISM: S. pneumoniae
29 <220> FEATURE:
32 <400> SEQUENCE: 1
33  atg aaa ttt agt aaa aaa tat ata gca gct gga tca gct gtt atc gta      48
34  tcc ttg agt cta tgt gcc tat gca cta aac cag cat cgt tcg cag gaa      96
35  aat aag gac aat aat cgt gtc tct tat gtg gat ggc agc cag tca agt     144
36  cag aaa agt gaa aac ttg aca cca gac cag gtt agc cag aaa gaa gga     192
37  att cag gct gag caa att gta atc aaa att aca gat cag ggc tat gta     240
38  acg tca cac ggt gac cac tat cat tac tat aat ggg aaa gtt cct tat     288
39  gat gcc ctc ttt agt gaa gaa ctc ttg atg aag gat cca aac tat caa     336
40  ctt aaa gac gct gat att gtc aat gaa gtc aag ggt ggt tat atc atc     384
41  aag gtc gat gga aaa tat tat gtc tac ctg aaa gat gca gct cat gct     432
42  gat aat gtt cga act aaa gat gaa atc aat cgt caa aaa caa gaa cat     480
43  gtc aaa gat aat gag aag gtt aac tct aat gtt gct gta gca agg tct     528
44  cag gga cga tat acg aca aat gat ggt tat gtc ttt aat cca gct gat     576
45  att atc gaa gat acg ggt aat gct tat atc gtt cct cat gga ggt cac     624
46  tat cac tac att ccc aaa agc gat tta tct gct agt gaa tta gca gca     672
47  gct aaa gca cat ctg gct gga aaa aat atg caa ccg agt cag tta agc     720
48  tat tct tca aca gct agt gac aat aac acg caa tct gta gca aaa gga     768
49  tca act agc aag cca gca aat aaa tct gaa aat ctc cag agt ctt ttg     816
50  aag gaa ctc tat gat tca cct agc gcc caa cgt tac agt gaa tca gat     864
51  ggc ctg gtc ttt gac cct gct aag att atc agt cgt aca cca aat gga     912
52  gtt gcg att ccg cat ggc gac cat tac cac ttt att cct tac agc aag     960
53  ctt tct gct tta gaa gaa aag att gcc aga atg gtg cct atc agt gga    1008
54  act ggt tct aca gtt tct aca aat gca aaa cct aat gaa gta gtg tct    1056
55  agt cta ggc agt ctt tca agc aat cct tct tct tta acg aca agt aag    1104
56  gag ctc tct tca gca tct gat ggt tat att ttt aat cca aaa gat atc    1152
57  gtt gaa gaa acg gct aca gct tat att gta aga cat ggt gat cat ttc    1200
58  cat tac att cca aaa tca aat caa att ggg caa ccg act ctt cca aac    1248

```

## RAW SEQUENCE LISTING

DATE: 10/09/2001

PATENT APPLICATION: US/09/471,255

TIME: 18:08:44

Input Set : A:\PTO.amc

Output Set: N:\CRF3\10092001\I471255.raw

```

59 aat agt cta gca aca cct tct cca tct ctt cca atc aat cca gga act 1296
60 tca cat gag aaa cat gaa gaa gat gga tac gga ttt gat gct aat cgt 1344
61 att atc gct gaa gat gaa tca ggt ttt gtc atg agt cac gga gac cac 1392
62 aat cat tat ttc ttc aag aag gac ttg aca gaa gag caa att aag gct 1440
63 gcg caa aaa cat tta gag gaa gtt aaa act agt cat aat gga tta gat 1488
64 tct ttg tca tct cat gaa cag gat tat cca ggt aat gcc aaa gaa atg 1536
65 aaa gat tta gat aaa aaa atc gaa gaa aaa att gct ggc att atg aaa 1584
66 caa tat ggt gtc aaa cgt gaa agt att gtc gtg aat aaa gaa aaa aat 1632
67 gcg att att tat ccg cat gga gat cac cat cat gca gat ccg att gat 1680
68 gaa cat aaa ccg gtt gga att ggt cat tct cac agt aac tat gaa ctg 1728
69 ttt aaa ccc gaa gaa gga gtt gct aaa aaa gaa ggg aat aaa gtt tat 1776
70 act gga gaa gaa tta acg aat gtt gtt aat ttg tta aaa aat agt acg 1824
71 ttt aat aat caa aac ttt act cta gcc aat ggt caa aaa cgc gtt tct 1872
72 ttt agt ttt ccg cct gaa ttg gag aaa aaa tta ggt atc aat atg cta 1920
73 gta aaa tta ata aca cca gat gga aaa gta ttg gag aaa gta tct ggt 1968
74 aaa gta ttt gga gaa gga gta ggg aat att gca aac ttt gaa tta gat 2016
75 caa cct tat tta cca gga caa aca ttt aag tat act atc gct tca aaa 2064
76 gat tat cca gaa gta agt tat gat ggt aca ttt aca gtt cca acc tct 2112
77 tta gct tac aaa atg gcc agt caa acg att ttc tat cct ttc cat gca 2160
78 ggg gat act tat tta aga gtg aac cct caa ttt gca gtg cct aaa gga 2208
79 act gat gct tta gtc aga gtg ttt gat gaa ttt cat gga aat gct tat 2256
80 tta gaa aat aac tat aaa gtt ggt gaa atc aaa tta ccg att ccg aaa 2304
81 tta aac caa gga aca acc aga acg gcc gga aat aaa att cct gta acc 2352
82 ttc atg gca aat gct tat ttg gac aat caa tcg act tat att gtg gaa 2400
83 gta cct atc ttg gaa aaa gaa aat caa act gat aaa cca agt att cta 2448
84 cca caa ttt aaa agg aat aaa gca caa gaa aac tca aaa ctt gat gaa 2496
85 aag gta gaa gaa cca aag act agt gag aag gta gaa aaa gaa aaa ctt 2544
86 tct gaa act ggg aat agt act agt aat tca acg tta gaa gaa gtt cct 2592
87 aca gtg gat cct gta caa gaa aaa gta gca aaa ttt gct gaa agt tat 2640
88 ggg atg aag cta gaa aat gtc ttg ttt aat atg gac gga aca att gaa 2688
89 tta tat tta cca tca gga gaa gtc att aaa aag aat atg gca gat ttt 2736
90 aca gga gaa gca cct caa gga aat ggt gaa aat aaa cca tct gaa aat 2784
91 gga aaa gta tct act gga aca gtt gag aac caa cca aca gaa aat aaa 2832
92 cca gca gat tct tta cca gag gca cca aac gaa aaa cct gta aaa cca 2880
93 gaa aac tca acg gat aat gga atg ttg aat cca gaa ggg aat gtg ggg 2928
94 agt gac cct atg tta gat cca gca tta gag gaa gct cca gca gta gat 2976
95 cct gta caa gaa aaa tta gaa aaa ttt aca gct agt tac gga tta ggc 3024
96 tta gat agt gtt ata ttc aat atg gat gga acg att gaa tta aga ttg 3072
97 cca agt gga gaa gtg ata aaa aag aat tta tct gat ttc ata gcg 3117
98 taa 3120

```

100 &lt;210&gt; SEQ ID NO: 2

101 &lt;211&gt; LENGTH: 1039

102 &lt;212&gt; TYPE: PRT

103 &lt;213&gt; ORGANISM: S. pneumoniae

105 &lt;400&gt; SEQUENCE: 2

106 Met Lys Phe Ser Lys Lys Tyr Ile Ala Ala Gly Ser Ala Val Ile Val

107 1 5 10 15

108 Ser Leu Ser Leu Cys Ala Tyr Ala Leu Asn Gln His Arg Ser Gln Glu

109 20 25 30

## RAW SEQUENCE LISTING

DATE: 10/09/2001

PATENT APPLICATION: US/09/471,255

TIME: 18:08:44

Input Set : A:\PTO.amc

Output Set: N:\CRF3\10092001\I471255.raw

```

110 Asn Lys Asp Asn Asn Arg Val Ser Tyr Val Asp Gly Ser Gln Ser Ser
111      35      40      45
112 Gln Lys Ser Glu Asn Leu Thr Pro Asp Gln Val Ser Gln Lys Glu Gly
113      50      55      60
114 Ile Gln Ala Glu Gln Ile Val Ile Lys Ile Thr Asp Gln Gly Tyr Val
115      65      70      75      80
116 Thr Ser His Gly Asp His Tyr His Tyr Tyr Asn Gly Lys Val Pro Tyr
117      85      90      95
118 Asp Ala Leu Phe Ser Glu Glu Leu Leu Met Lys Asp Pro Asn Tyr Gln
119      100      105      110
120 Leu Lys Asp Ala Asp Ile Val Asn Glu Val Lys Gly Gly Tyr Ile Ile
121      115      120      125
122 Lys Val Asp Gly Lys Tyr Tyr Val Tyr Leu Lys Asp Ala Ala His Ala
123      130      135      140
124 Asp Asn Val Arg Thr Lys Asp Glu Ile Asn Arg Gln Lys Gln Glu His
125      145      150      155      160
126 Val Lys Asp Asn Glu Lys Val Asn Ser Asn Val Ala Val Ala Arg Ser
127      165      170      175
128 Gln Gly Arg Tyr Thr Thr Asn Asp Gly Tyr Val Phe Asn Pro Ala Asp
129      180      185      190
130 Ile Ile Glu Asp Thr Gly Asn Ala Tyr Ile Val Pro His Gly Gly His
131      195      200      205
132 Tyr His Tyr Ile Pro Lys Ser Asp Leu Ser Ala Ser Glu Leu Ala Ala
133      210      215      220
134 Ala Lys Ala His Leu Ala Gly Lys Asn Met Gln Pro Ser Gln Leu Ser
135      225      230      235      240
136 Tyr Ser Ser Thr Ala Ser Asp Asn Asn Thr Gln Ser Val Ala Lys Gly
137      245      250      255
138 Ser Thr Ser Lys Pro Ala Asn Lys Ser Glu Asn Leu Gln Ser Leu Leu
139      260      265      270
140 Lys Glu Leu Tyr Asp Ser Pro Ser Ala Gln Arg Tyr Ser Glu Ser Asp
141      275      280      285
142 Gly Leu Val Phe Asp Pro Ala Lys Ile Ile Ser Arg Thr Pro Asn Gly
143      290      295      300
144 Val Ala Ile Pro His Gly Asp His Tyr His Phe Ile Pro Tyr Ser Lys
145      305      310      315      320
146 Leu Ser Ala Leu Glu Glu Lys Ile Ala Arg Met Val Pro Ile Ser Gly
147      325      330      335
148 Thr Gly Ser Thr Val Ser Thr Asn Ala Lys Pro Asn Glu Val Val Ser
149      340      345      350
150 Ser Leu Gly Ser Leu Ser Ser Asn Pro Ser Ser Leu Thr Thr Ser Lys
151      355      360      365
152 Glu Leu Ser Ser Ala Ser Asp Gly Tyr Ile Phe Asn Pro Lys Asp Ile
153      370      375      380
154 Val Glu Glu Thr Ala Thr Ala Tyr Ile Val Arg His Gly Asp His Phe
155      385      390      395      400
156 His Tyr Ile Pro Lys Ser Asn Gln Ile Gly Gln Pro Thr Leu Pro Asn
157      405      410      415
158 Asn Ser Leu Ala Thr Pro Ser Pro Ser Leu Pro Ile Asn Pro Gly Thr

```

## RAW SEQUENCE LISTING

DATE: 10/09/2001

PATENT APPLICATION: US/09/471,255

TIME: 18:08:44

Input Set : A:\PTO.amc

Output Set: N:\CRF3\10092001\I471255.raw

159		420		425		430	
160	Ser His Glu Lys His Glu Glu Asp Gly Tyr Gly Phe Asp Ala Asn Arg						
161		435		440		445	
162	Ile Ile Ala Glu Asp Glu Ser Gly Phe Val Met Ser His Gly Asp His						
163		450		455		460	
164	Asn His Tyr Phe Phe Lys Lys Asp Leu Thr Glu Glu Gln Ile Lys Ala						
165	465		470		475		480
166	Ala Gln Lys His Leu Glu Glu Val Lys Thr Ser His Asn Gly Leu Asp						
167		485		490		495	
168	Ser Leu Ser Ser His Glu Gln Asp Tyr Pro Gly Asn Ala Lys Glu Met						
169		500		505		510	
170	Lys Asp Leu Asp Lys Lys Ile Glu Glu Lys Ile Ala Gly Ile Met Lys						
171		515		520		525	
172	Gln Tyr Gly Val Lys Arg Glu Ser Ile Val Val Asn Lys Glu Lys Asn						
173		530		535		540	
174	Ala Ile Ile Tyr Pro His Gly Asp His His His Ala Asp Pro Ile Asp						
175	545		550		555		560
176	Glu His Lys Pro Val Gly Ile Gly His Ser His Ser Asn Tyr Glu Leu						
177		565		570		575	
178	Phe Lys Pro Glu Glu Gly Val Ala Lys Lys Glu Gly Asn Lys Val Tyr						
179		580		585		590	
180	Thr Gly Glu Glu Leu Thr Asn Val Val Asn Leu Leu Lys Asn Ser Thr						
181		595		600		605	
182	Phe Asn Asn Gln Asn Phe Thr Leu Ala Asn Gly Gln Lys Arg Val Ser						
183		610		615		620	
184	Phe Ser Phe Pro Pro Glu Leu Glu Lys Lys Leu Gly Ile Asn Met Leu						
185	625		630		635		640
186	Val Lys Leu Ile Thr Pro Asp Gly Lys Val Leu Glu Lys Val Ser Gly						
187		645		650		655	
188	Lys Val Phe Gly Glu Gly Val Gly Asn Ile Ala Asn Phe Glu Leu Asp						
189		660		665		670	
190	Gln Pro Tyr Leu Pro Gly Gln Thr Phe Lys Tyr Thr Ile Ala Ser Lys						
191		675		680		685	
192	Asp Tyr Pro Glu Val Ser Tyr Asp Gly Thr Phe Thr Val Pro Thr Ser						
193		690		695		700	
194	Leu Ala Tyr Lys Met Ala Ser Gln Thr Ile Phe Tyr Pro Phe His Ala						
195	705		710		715		720
196	Gly Asp Thr Tyr Leu Arg Val Asn Pro Gln Phe Ala Val Pro Lys Gly						
197		725		730		735	
198	Thr Asp Ala Leu Val Arg Val Phe Asp Glu Phe His Gly Asn Ala Tyr						
199		740		745		750	
200	Leu Glu Asn Asn Tyr Lys Val Gly Glu Ile Lys Leu Pro Ile Pro Lys						
201		755		760		765	
202	Leu Asn Gln Gly Thr Thr Arg Thr Ala Gly Asn Lys Ile Pro Val Thr						
203		770		775		780	
204	Phe Met Ala Asn Ala Tyr Leu Asp Asn Gln Ser Thr Tyr Ile Val Glu						
205	785		790		795		800
206	Val Pro Ile Leu Glu Lys Glu Asn Gln Thr Asp Lys Pro Ser Ile Leu						
207		805		810		815	

## RAW SEQUENCE LISTING

DATE: 10/09/2001

PATENT APPLICATION: US/09/471,255

TIME: 18:08:44

Input Set : A:\PTO.amc

Output Set: N:\CRF3\10092001\I471255.raw

```

208 Pro Gln Phe Lys Arg Asn Lys Ala Gln Glu Asn Ser Lys Leu Asp Glu
209           820           825           830
210 Lys Val Glu Glu Pro Lys Thr Ser Glu Lys Val Glu Lys Glu Lys Leu
211           835           840           845
212 Ser Glu Thr Gly Asn Ser Thr Ser Asn Ser Thr Leu Glu Glu Val Pro
213           850           855           860
214 Thr Val Asp Pro Val Gln Glu Lys Val Ala Lys Phe Ala Glu Ser Tyr
215           865           870           875           880
216 Gly Met Lys Leu Glu Asn Val Leu Phe Asn Met Asp Gly Thr Ile Glu
217           885           890           895
218 Leu Tyr Leu Pro Ser Gly Glu Val Ile Lys Lys Asn Met Ala Asp Phe
219           900           905           910
220 Thr Gly Glu Ala Pro Gln Gly Asn Gly Glu Asn Lys Pro Ser Glu Asn
221           915           920           925
222 Gly Lys Val Ser Thr Gly Thr Val Glu Asn Gln Pro Thr Glu Asn Lys
223           930           935           940
224 Pro Ala Asp Ser Leu Pro Glu Ala Pro Asn Glu Lys Pro Val Lys Pro
225           945           950           955           960
226 Glu Asn Ser Thr Asp Asn Gly Met Leu Asn Pro Glu Gly Asn Val Gly
227           965           970           975
228 Ser Asp Pro Met Leu Asp Pro Ala Leu Glu Glu Ala Pro Ala Val Asp
229           980           985           990
230 Pro Val Gln Glu Lys Leu Glu Lys Phe Thr Ala Ser Tyr Gly Leu Gly
231           995          1000          1005
232 Leu Asp Ser Val Ile Phe Asn Met Asp Gly Thr Ile Glu Leu Arg Leu
233           1010          1015          1020
234 Pro Ser Gly Glu Val Ile Lys Lys Asn Leu Ser Asp Phe Ile Ala
235           1025          1030          1035
237 <210> SEQ ID NO: 3
238 <211> LENGTH: 2523
239 <212> TYPE: DNA
240 <213> ORGANISM: S. pneumoniae
242 <220> FEATURE:
243 <221> NAME/KEY: CDS
244 <222> LOCATION: (1)...(2520)
245 <223> OTHER INFORMATION: Coding region of BVH-11 gene
247 <400> SEQUENCE: 3
248 atg aaa atc aat aaa aaa tat cta gct ggg tca gta gct aca ctt gtt      48
249 Met Lys Ile Asn Lys Lys Tyr Leu Ala Gly Ser Val Ala Thr Leu Val
250   1           5           10           15
252 tta agt gtc tgt gct tat gaa cta ggt ttg cat caa gct caa act gta      96
253 Leu Ser Val Cys Ala Tyr Glu Leu Gly Leu His Gln Ala Gln Thr Val
254           20           25           30
256 aaa gaa aat aat cgt gtt tcc tat ata gat gga aaa caa gcg acg caa      144
257 Lys Glu Asn Asn Arg Val Ser Tyr Ile Asp Gly Lys Gln Ala Thr Gln
258           35           40           45
260 aaa acg gag aat ttg act cct gat gag gtt agc aag cgt gaa gga atc      192
261 Lys Thr Glu Asn Leu Thr Pro Asp Glu Val Ser Lys Arg Glu Gly Ile
262           50           55           60

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/471,255

DATE: 10/09/2001

TIME: 18:08:45

Input Set : A:\PTO.amc

Output Set: N:\CRF3\10092001\I471255.raw

L:17 M:270 C: Current Application Number differs, Replaced Current Application No

L:17 M:271 C: Current Filing Date differs, Replaced Current Filing Date